DOCUMENT RESUME

ED 251 470 TM 840 491

AUTHOR Matter, M. Kevin; Ligon, Glynn
TITLE Issues in Developing a Local Minimum Competency Test: Our BEST Effort.
INSTITUTION Austin Independent School District, Tex. Office of Research and Evaluation.
REPORT NO AISD-ORE-83.44
PUB DATE Apr 84
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
AVAILABLE FROM Office of Research and Evaluation, AISD, 6100 Guadalupe, Box 79, Austin, TX, 78752.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Basic Skills; Board of Education Policy; *Graduation Requirements; *High Schools; Item Analysis; *Item Banks; Mathematics Skills; *Minimum Competency Testing; Reading Skills; Remedial Instruction; *Test Construction; Testing Programs
IDENTIFIERS *Austin Independent School District TX; *Basic Educational Skills Test; Iowa Tests of Basic Skills; Test Security

ABSTRACT

This paper describes the development of the Basic Educational Skills Test (BEST) in the Austin Independent School District. In 1982 future options for the AISD minimum competency program, which requires all students to meet minimum proficiency levels in reading and mathematics before graduation, were evaluated against legal and local criteria. It was decided to develop an item bank that would allow the generation of multiple, parallel forms of a competency test to be administered: to students in tutorial courses; to senior transfers; and in special testing sessions to students meeting the course requirements, but not yet meeting the minimum competency requirement, for graduation. Important issues during the BEST development process fell into six general categories: (1) identification of objectives; (2) item acquisition; (3) item calibration; (4) test formatting; (5) test production; and (6) item and test booklet security. (BS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED251470

Issues in Developing a Local Minimum Competency Test: Our BEST Effort

M. Kevin Matter

Glynn Ligon

Austin Independent School District

Office of Research and Evaluation

Austin, Texas

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

K. Matter

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

A Paper Presented at the Annual Meeting of the American Educational Research Association, New Orleans, April, 1984

ORE Publication Number 83.44

491
C48
M47

In the mid-1970's, the Austin Independent School District (AISD) joined the trend by adopting minimum competency requirements for high school graduation. Today, we have abandoned the original tests, developed our own (Austin's BEST), raised our standards, and are dealing with pressure to raise standards even more. In this process, we have met head-on the major and minor issues incumbent in any minimum competency testing effort. Let us share some of our experience with you.

A BRIEF HISTORY OF AISD'S MINIMUM COMPETENCY REQUIREMENTS

The AISD School Board approved a minimum competency requirement for high school graduation in 1975, revised in 1976 to require an eighth-grade level of proficiency in reading and mathematics. Students could meet the competency requirements in the eighth grade on the California Achievement Tests (CAT), Level 4, Form A or B, or in grades 9-12 on the Sequential Tests of Educational Progress (STEP), Series II, Form A or B.

Three events occurred during the 1979-80 school year which affected the minimum competency requirements.

* The District replaced the CAT with the Iowa Tests of Basic Skills (ITBS), Form 7.
* That same year the Board raised the level required for graduation from an eighth-grade level (8.5 GE) to the ninth-grade level (9.0 GE), beginning with May, 1983 graduates.
* The Texas Assessment of Basic Skills (TABS), a state-mandated, criterion-referenced test, was given to all students in grades 3, 5, and 9. We equated the TABS with the 9.0 GE on the STEP to enable students to meet AISD competency requirements on the TABS (Carsrud and Ligon, 1981).

The AISD minimum competency program incorporates several features that attempt to ensure that all students meet the minimum levels of proficiency in reading and mathematics by the time they graduate.

* Students may meet the requirements as early as the eighth grade, and have yearly opportunities to do so through the annual achievement testing.
* Tutorial courses are offered for students who have not met the requirements by the time they have completed the basic reading and mathematics courses required for graduation.

* Special testing sessions are provided for senior transfers and other students in grades 10-12, to assure that students who need special help are placed in tutorial courses.
* Several tests are available on which to meet the requirements. The current tests given are the ITBS, TABS, the Tests of Achievement and Proficiency (TAP), and Austin's Basic Educational Skills Test (BEST).

A YEAR OF DECISION

During the 1981-82 school year, we outlined the options that were available for the future. These included:

1. Continuing the program as it currently existed.
2. Using the statewide TABS alone as the competency measure.
3. Using the annual standardized achievement test administrations as the only competency measures.
4. Developing an item bank that would allow the generation of multiple, parallel forms of a locally developed competency test.
5. Incorporating a locally developed competency test and item bank into the current program. This would combine the desirable features of options 1 and 4.

To evaluate these options, we examined the six characteristics of a legally defensible competency program, as defined by recent court cases (Madaus, 1983).

1. Valid objectives which describe those skills which are truly basic competencies
2. A valid measure of those objectives
3. Assurance that the skills are actually taught
4. Early assessment and identification of those needing remediation
5. Remedial or tutorial assistance for those needing it
6. Multiple opportunities to pass the competency test

A comparison of the five options for AISD on the six characteristics is shown in Figure 1.

In the spring of 1982, the decision was made to develop an item bank that would allow the generation of multiple, parallel forms of a competency test to be administered to students in tutorial courses, to senior transfers, and in the special testing sessions provided for students who had met course requirements for graduation but had not yet met the minimum competency requirements.

Our task in the development of the test item bank was to address both the legal needs outlined above and local needs regarding acceptance of the item bank concept by administrators and teachers. The issues which became important during the development process fell into six general categories.

COMPETENCY PROGRAM CHARACTERISTICS	TYPE OF PROGRAM				Comments
	Current Program	TABS Only	Standardized Tests Only	Item Bank	
Valid Objectives	?	+	?	+	An item bank built around the TABS objectives uses objectives set through an elaborate statewide effort. Current program objectives were shaped by the CAT skills rather than being set from the ground up.
Valid Measure	?	+	?	+	All tests can probably be shown to be valid for the objectives/skills measured. However, unless the objectives are valid, the test cannot be.
Skills Actually Taught	-	+	-	+	Standardized tests measure such a wide range of skills that assuring that all are taught is problematic.
Early Assessment	+	+	+	+	All begin by grade 9.
Remedial/Tutorial Assistance	+	+	+	+	All identify students prior to start of tutorials in grade 11.
Multiple Opportunities	+	?	?	+	TABS and standardized tests allow testing just once per year.

+ = Strength ? = Unknown - = Weakness

Figure 1. Comparison of strengths and weaknesses of competency programs on six characteristics of a legally defensible program.

a) Identification of Objectives

b) Acquisition of Items

c) Calibration of Items

d) Formatting of Tests

e) Production of Tests

f) Security of Items and Test Booklets

Identification of Objectives

To make our competency program more legally defensible, the tutorial course objectives were revised during the summer of 1982. The statewide TABS objectives were adopted as the basic competencies required for graduation from AISD. The TABS objectives had been set through an elaborate statewide effort, so they should meet the legal definition of valid objectives.

Acquisition of Items

One of our goals in the item bank development process was to avoid traditional pilot testing to conserve student instructional time. We wanted to pilot test students and be able to use the results of that testing as an opportunity to meet the AISD graduation requirements. At that time we were giving the ITBS, Level 14, as the final exam in the tutorial courses. The

strategy we adopted centered around leasing particular ITBS items from The Riverside Publishing Company. The leased items enabled us to assess competency for those students tested since we knew the difficulty level of those items. At the same time, we could calibrate our research items with the leased ITBS items, eliminating separate pilot testing of items for that purpose.

Our timeline was short (April to November) for development of our item bank and multiple forms of a test incorporating leased ITBS items and research items. In that time, period we needed to locate and purchase items from existing item banks, select usable items, write additional items as needed, and select the ITBS items which should be used to assess competency and calibrate the research items.

We discovered quickly that few available item banks contained very many items at a ninth-grade level in reading and mathematics. We purchased items that we could, including some from the states of California and Michigan, the Northwest Educational Research Laboratory, and the Education Commission of the States (National Assessment of Educational Progress).

Tutorial teachers were contracted to review the purchased items for item difficulty and match with the new tutorial course objectives. We wanted to develop at least five forms of the test in both reading and mathematics using purchased items for the fall, 1982 testings. Few of the purchased items were usable without modification, and by the time the forms were completed,

almost all usable purchased items had been incorporated into our fall, 1982 test forms. That meant that future forms would require new items to be written from scratch.

Calibration of the Items

For the 1982-83 school year, all of the items which came from purchased item banks and those written by teachers and staff members were research items. Competency status would be assessed only by ITBS items of known difficulty. While research items were being reviewed and written, ITBS Level 14 items were selected which were around the 9.0 difficulty level. Using past years' districtwide test results, we Rasch calibrated the ITBS Level 14 reading and mathematics items. A constraint in selecting the reading items was the need to select groups of items which pertained to one reading selection. The ITBS was normed with a particular number of items for each reading selection. We needed to use a selection and all associated test items to maintain the integrity of each individual item. Thus, reading selections were chosen on the basis of having their group test item average as close to 9.0 GE as possible. We leased 80 items (38 reading, 42 mathematics) for adequate reliability and validity.

The Rasch scaling process (one parameter) was used to calibrate the research items with the leased ITBS items.

During the 1982-83 school year, the leased ITBS items were used to measure competency on Austin's BEST. The ITBS and research items were grouped by pages, with all ITBS items on one page, all research items on the next page. During the summer of 1983, the research items were calibrated with the ITBS items.

Formatting of Tests

One of our major concerns was that Austin's BEST look professionally made. We wanted students, teachers, instructional coordinators, the Superintendent, and other administrative staff to feel that Austin's BEST was made as well as, or better than, the standardized tests used in the District. Therefore, the design and format of Austin's BEST was critical for acceptance by these groups.

We solicited input from tutorial teachers and central office instructional staff on several aspects of Austin's BEST. The factors considered in each area are outlined below.

TEST LENGTH:

. The test should be long enough to be reliable, but limiting noninstructional testing time as much as possible.

LENGTH OF ITEMS:

. The test should consist of brief reading selections, to reflect the shorter attention span and higher frustration level of many students at this achievement level.

TEST TIMING:

. There should be a preference for a power test over a speed test since the purpose is to measure ability to master particular objectives.

NOTATION AND STYLE:

. Mathematical notation and style should be consistent with that used in the tutorial courses and AISD curriculum. A guide for typing Austin's BEST mathematics items was developed to provide consistency across test forms (Attachment A).

ITEM SEQUENCING:

. Tutorial teachers felt that mathematical computation problems should be placed at the beginning of the test since their students are more proficient in those skills than in problem solving or concepts. The ITBS placed computation problems last on the test, so we needed to determine if item placement on the test affected item calibration. We developed two Austin's BEST forms with the only difference being location of computation and concepts research items. In these two BEST forms, the ITBS items maintained their "concepts items first -- computation items last" format, but

the computation research items followed the ITBS concepts items on one form and followed the ITBS computation items on the other form (see Figure 2). Item location did influence calibration of identical items. Since we wanted computation items in the beginning of the test, we used the calibrations obtained from having them follow the ITBS concepts items in the first third of the test.

	FORM 1	FORM 2
Page 1	ITBS Concepts	ITBS Concepts
Page 2	Research Concepts	Research Computation
Page 3	ITBS Problems	ITBS Problems
Page 4	Research Problems	Research Problems
Page 5	ITBS Computation	ITBS Computation
Page 6	Research Computation	Research Concepts

Figure 2.

ANSWER SHEETS:

. Answer sheets should be custom made for Austin's BEST, so students will not have difficulty in marking their answers. Although we preferred a custom answer sheet, in 1981-82 we used a ready-made answer sheet to allow flexibility and

because our short timeline prohibited its development. This ready-made answer sheet had two drawbacks that we needed to address: extra items and an extra choice for each item. We worked around the problem of extra items on the sheet by printing a message to "stop here" on each answer sheet after the final legitimate test item number. Students were also told that option "E" was never a correct choice and that no mark should be made in that spot or the item would be scored as incorrect because of double-bubbling. (Some teachers had more problems with this than the students did. Several times teachers told their students prior to testing to mark through all the E options right away so they would not mark it accidentally during the test!) A custom answer sheet was developed for the 1983-84 school year to eliminate these problems.

Production of Tests

Production and reproduction of Austin's BEST were important in making a professional appearance. Typesetting the reading forms was needed to get as much as possible on a page while maintaining an adequate amount of white space. Typesetting reduced the number of pages, lowering the cost of reproducing the test. We would have preferred to typeset the mathematics test, but a type font which produces fractions with a horizontal line separating the

numerator and denominator (instead of a diagonal line) was not available. Typing the mathematics items was our only alternative.

It took two weeks to typeset, proof, and correct the reading tests. The mathematics tests took three typings before the format and spacing were acceptable, then fractions, lines, charts, tables, and other final touches were added. In all, two weeks were needed to finalize five mathematics forms.

The color and weight of paper for reproduction were carefully considered, also. In order to identify the reading tests from the mathematics, two colors of paper were used, an ivory and a light tan. The colors chosen would print well and not promote eye strain, yet be more attractive than white paper.

Security of Items and Test Booklets

The printing was done commercially, with security of the tests stressed. All negatives, printing plates, samples, and other waste from the printing process were returned to our office for secure disposal.

All copies of Austin's BEST were stamped with a sequential number as an aid in keeping track of the copies as they were used for test administrations.

References

Carsrud, K. B., & Ligon, G. Equating studies: A manual of issues, options, and decisions for public school evaluators. Paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles, April, 1981. (ORE Publication Number 80.50)

Madaus, G. F. The courts, validity, and minimum competency testing. Boston: Kluwer-Nijhoff Publishing, 1983.